Three Ways in Which Pandemic Models May Perform a Pandemic

PHILIPPE VAN BASSHUYSEN
Leibniz University Hannover

LUCIE WHITE
Leibniz University Hannover

DONAL KHOSROWI
Leibniz University Hannover

MATHIAS FRISCH
Leibniz University Hannover

Abstract: Models not only represent but may also *influence* their targets in important ways. While models' abilities to influence outcomes has been studied in the context of economic models, often under the label 'performativity', we argue that this phenomenon also pertains to epidemiological models, such as those used for forecasting the trajectory of the Covid-19 pandemic. After identifying three ways in which a model by the Covid-19 Response Team at Imperial College London (Ferguson et al. 2020) may have influenced scientific advice, policy, and individual responses, we consider the implications of epidemiological models' performative capacities. We argue, first, that performativity may impair models' ability to successfully predict the course of an epidemic; but second, that it may provide an additional sense in which these models can be successful, namely by changing the course of an epidemic.

Keywords: Covid-19, epidemiological models, performativity, prediction, success, model evaluation

JEL Classification: B40, C63, I12, I18, Z18

I. INTRODUCTION

Philosophers are increasingly scrutinising epidemiological models that have been used to forecast the trajectory of the Covid-19 pandemic (Horner and Symons 2020; Nguyen and Frigg 2021; Winsberg, Brennan, and Surprenant 2020; Schroeder 2021; van Basshuysen and White 2021). The success of these models is often evaluated in terms of their predictive accuracy (Friedman et al. 2021). We argue that there is an additional dimension along which epidemiological models can be successful, namely in *influencing* the course of an epidemic. Our argument draws on the concept of 'performativity', which economic sociologists and philosophers of economics have used to highlight the fact that economic models may not only represent parts of the economy, but can also influence the behaviours of the economic systems modelled (Callon 1998; MacKenzie 2006; Guala 2007; Mäki 2013; Boldyrev and Ushakov 2016). This notion has a long history in philosophy of science (see Buck 1963), economics, and the social sciences (see Henshel 1993), where it is often referred to as 'reflexive prediction' (Buck 1963; Henshel 1993).[1] Predictions (or 'prophecies') can be 'self-fulfilling' or 'self-defeating', understood broadly to mean that the dissemination of the prediction is causally relevant in either bringing about the event predicted or in preventing it from occurring (Buck 1963). We similarly understand the notion of performativity in a broad way; models *perform* by changing important properties of the phenomena they are representing (see Mäki 2013).[2] Whereas earlier accounts tended to focus on the negative aspects of this phenomenon (particularly the difficulties it poses for accurate prediction), contemporary work on performativity also emphasises potential positive aspects, particularly how an understanding of performativity can allow us to intervene constructively in the economy (Guala 2007).

While there has been sustained discussion of the performativity of economic models, the question of whether epidemiological models might be performative has thus far been neglected.[3] We aim to make a first step

---

[1] Ian Hacking (1999) describes a related phenomenon involving social kinds.

[2] Mäki himself is critical of attempts to characterize this type of phenomenon in the context of economic modelling in terms of Austin's notion of performativity, arguing that Austin's notion is concerned with acts being *constituted* by certain utterances. By contrast, according to Mäki, economic theorizing may have *causal* consequences for economic activity without being constitutive of it. We set aside the question of whether Austin's notion is apt in this context here; rather, we wish simply to draw attention to the fact that, like economic models, epidemiological models can have causal effects on the systems modelled.

[3] Some social scientists have described epidemiological models as performative (Rhodes

towards closing this gap by arguing, first, that epidemiological models can be performative in several ways, and second, that this has important implications for how these models should be regarded and evaluated. We proceed as follows: after sketching some important features of epidemiological models, we focus on a model that was used by researchers at Imperial College London (ICL) to forecast the pandemic in the UK and the US (Ferguson et al. 2020). On the basis of this example, and focused particularly on its use in the UK, we identify three ways in which models may be performative: (i) models may influence policy advising; (ii) models may affect policy-makers' responses to an epidemic; and (iii) models may directly influence individual responses to an epidemic. While (i) and (ii) are well-documented in the case of the ICL model, (iii) is more speculative.[4] Having identified the ways in which epidemiological models can be performative, we elaborate upon the ways in which their performative aspects may complicate the task of successfully forecasting the course of an epidemic. At the same time, however, we contend, their performative abilities provide an additional way in which these models can be successful, namely by changing the course of an epidemic.

II. Epidemiological Models

Epidemiological models are among the instruments that policy-makers have used (and continue to use) when making policy decisions during the Covid-19 pandemic. These models provide highly idealized representations of how the virus will spread through a population, under a variety of assumptions.

The simplest models are SIR or SEIR models, which assign members of a population to different subgroups, 'the *S*usceptible', '*I*nfectious', '*Re*moved/*R*ecovered', and, in the case of SEIR models, also the '*E*xposed'. SIR and SEIR models consist of several deterministic equations that describe how the proportional sizes of the different subgroups change with time based on transition rates between these groups. These rates are represented by a small number of parameters: the transmission rate, the recovery rate, and the latency period. SIR or SEIR models are macroscopic models that treat populations as a whole and abstract from interactions

---

et al. 2020), but their focus differs from ours, as their main concern is that we should adapt these models to the changing circumstances.

[4] The first two types of 'performativity' we outline could apply to any model that is used to inform policy, whereas the third sense is more distinctive and methodologically interesting, and it is on this notion that most theorists focus (see, for example, the articles in MacKenzie, Muniesa, and Siu 2007).

among individuals and transmission events. The model used by the Imperial College London to model the global evolution of the pandemic in different regions of the world is a SEIR model (Walker et al. 2020).

By contrast, the 'CovidSim' model discussed in the highly influential *Report 9* by the Imperial College Covid-19 Response Team, upon which we will chiefly focus here, is an individual-based model (Ferguson et al. 2020), in which individuals are assigned to different types of location where contacts occur—within the household, at school, in the workplace, and in the wider community. CovidSim models transmission events through contacts among individuals at these locations and movements between locations to derive the evolution of infection numbers. CovidSim is a stochastic model containing a very large number of parameters, many of which are only poorly constrained by empirical data. But a sensitivity analysis (Edeling et al. 2021) has shown that the model is particularly sensitive to variations in three parameters: the length of the latent period in which a patient has no symptoms and is not infectious; the delay to start case isolation; and, the 'relative spatial contact rate given social distancing' parameter, which captures the effectiveness of social distancing (this last parameter will be particularly important in what follows).

III. Three Performative Aspects of the ICL Model

In the following, we will outline three salient ways in which epidemiological models (with a particular focus on the above-mentioned, much-discussed ICL model presented in *Report 9*) might be thought to have influenced the course of the Covid-19 pandemic.

III.I. The ICL Model Performed Scientific Advising

In a recent working paper, Jonathan Birch (2020) documents how the provision of scientific advice in the UK evolved in the lead up to the introduction of the ICL model used in *Report 9* (Ferguson et al. 2020). He suggests that the Scientific Advisory Group for Emergencies (SAGE)—the group chiefly responsible for guiding the UK's initial pandemic response—dispensed advice in three distinctive forms during March 2020. They began by providing advice in the form of what Birch (2020, 6) refers to as "no unconditional recommendations": that is, they presented a list of potentially effective means of pursuing a variety of goals that policymakers might have, with no indication of which goals might be desirable or worthy of pursuit. As the situation worsened, by March 9, SAGE shifted to providing advice in the form of "disjunctive unconditional

recommendations" (Birch 2020, 6); endorsing a specific end but leaving open a variety of means for achieving this end. For example, during this period, they outlined a range of non-pharmaceutical interventions that they believed could, in some combination, achieve epidemic mitigation, and recommended that some subset of these measures be implemented (SAGE 2020a). On March 18, shortly after SAGE had begun to discuss the ICL modelling paper, the form of advice shifted again, to a "single unconditional recommendation" (Birch 2020, 6), namely, that school closures were immediately required in order to prevent the NHS from becoming overwhelmed (SAGE 2020b).

Birch traces this trajectory in order to argue that the latter form of advice, while perhaps not appropriate in normal situations, is appropriate in fast-moving crisis situations. But he also comments in passing on a possible reason for this final shift to advice in the form of a single unconditional recommendation. Namely, two days before issuing their single unconditional recommendation (on March 16), the SAGE team discussed the ICL modelling paper for the first time. This paper displayed data about critical care capacity in the same graph as projections about demand for critical care in the absence of strict suppression measures (Figure 1), revealing, in stark form, the extent to which demand was predicted to outstrip supply, even with existing mitigation measures in place. Birch suggests that this might have been the first time that some SAGE members were made aware of the dramatic size of the mismatch. The ICL modellers also included a single unconditional recommendation in the report on the basis of their projections: "Epidemic suppression is the only viable strategy at the current time" (Ferguson et al. 2020, 16). It is possible, then, that this model—particularly through the graphical representation of its results[5]—was instrumental in shaping the form of the scientific advice provided by SAGE.

III.II. The Model Performed Policy-Makers' Responses to the Pandemic
It is clear that the ICL model had a significant impact on public policy in the UK; on March 16, immediately following the release of *Report 9*, the UK drastically altered (Bosely 2020) its previous, mitigation-based strategy (DHSC 2020), and almost all of the ICL team's recommended measures were implemented in the following days (see van Basshuysen

---

[5] The ways in which graphical representations of models and their results may influence how these models are adopted is an important topic in the performativity literature; see, for instance, Pahl and Sparsam (2016) in the context of the IS-LM model.
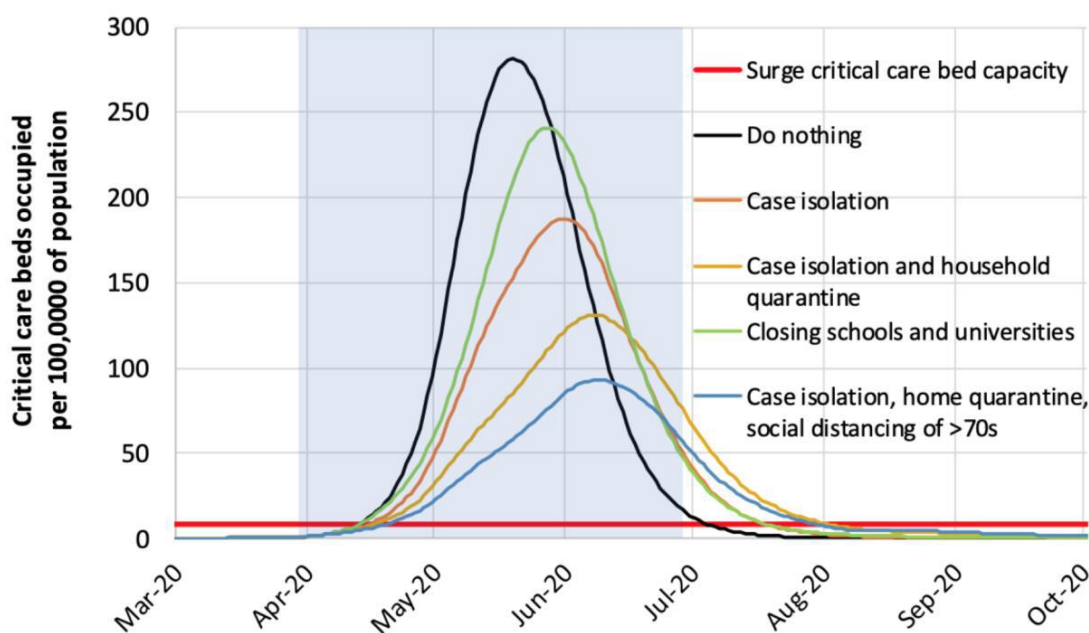
Figure 1: Projected demand for critical care beds under different possible mitigation-based policies is displayed together with surge critical care bed capacity. *Source:* Ferguson et al. (2020).

and White, forthcoming). This may have further led, however, to some noteworthy interactions between the epidemiological models that provided a basis for policy advice, and the responses of policy-makers. The ICL model and models like it, which project the impact of various public policy measures, are designed to provide policy guidance. Their usefulness depends on their predictive performance. But the implementation of the modellers' recommendations steers the course of the epidemic away from a model's reference forecasts. This can lead to claims that the model is 'too pessimistic'; that is, that its projections about what would happen in the absence of interventions or behaviour changes (but where mitigation measures are implemented) do not match the observed numbers of infected, hospitalized, and dead (see, for example, St. Onge and Campan 2020). We might regard this as a distinctive instance of *counterperformativity* (see Mackenzie 2006)—the implementation of the recommended policies steers the real-world outcome away from the reference forecast, but notably, in the case of the ICL model, this was the exact effect that the modellers were hoping to achieve.[6]

---

[6] Although this phenomenon has led to criticism of such models for generating inaccurate predictions, and comparison of reference forecasts to actual outcomes is used to evaluate the predictive accuracy of these models (see, for example, Friedman et al. 2021), it should be emphasised that the projection here is contingent on no measures being implemented (that is, this is a conditional prediction). It is helpful here to distinguish

Joseph Friedman et al. (2021) draw attention to another problem in their evaluation of the predictive performance of various international Covid-19 mortality forecasting models. They suggest that transmission-based epidemiological models tend to overestimate transmission in their reference forecasts, because prolonged exponential growth in transmission rarely continues unabated—governments (and/or individuals, see below) will change their behaviour in response to deteriorating circumstances—and, we contend, in response to updated modelling projections. The accuracy of modelling predictions and projections could be improved, they argue, by taking such endogenous factors into account (which, as they note, are often explicitly included in economic analyses). In addition, they suggest that models (such as the ICL models) that aim to predict the impact of various public policy measures would be better evaluated by looking at the error in the difference between the reference forecast and the forecast for the implemented policy. Adequately modelling endogenous responses may be complicated, however, by the third performative aspect of epidemiological models discussed below. Moreover, as Friedman et al. (2021) concede, the relevant counterfactuals can be difficult to assess. Even when comparing the projection concerning the implemented policy to its actual effects, the effects of the policy must be somehow distinguished from other factors. It should, however, be noted that economists have made progress in recognising and incorporating endogeneity into Covid-19 modelling, which, among other things, can assist in clarifying the role of government policy in disease dynamics (see Avery et al. 2020).

The belief that models were inaccurate and overly pessimistic might have also had potential further follow-on effects for the implementation of the policy advice on the basis of further models. It is possible that the scepticism and controversy surrounding the ICL projections was one of the factors that contributed to a reluctance to implement a further, short term, 'circuit-breaker' lockdown when a second wave was unfolding in the

---

predictions, which generate statements about future states of the actual world, from projections, which generate statements about counterfactual outcomes, conditional on some set of counterfactual assumptions. The performative effect of the model consists in influencing which member of a set of projections ends up most closely resembling the trajectory of the epidemic in the actual world and does not compromise its predictive accuracy. While the model steers the trajectory of the epidemic further away from the trajectory in the reference projection, this does not undermine the predictive accuracy of the model. Thus, this does not constitute the type of 'self-defeating prediction' that is the focus of much attention in the performativity literature, where, for example, the prediction that an imminent disaster *will happen* prevents this disaster from happening and thus undermines the accuracy of the prediction (see, for example, Henshel 1993).

UK in autumn 2020 (see Sayers 2020), which SAGE called for with increasing urgency (Stewart and Sample 2020) on the basis of mathematical modelling projections (SAGE 2020c). The refusal to implement this advice constituted a significant break between scientific advisors and policy-makers in the UK (Stewart and Sample 2020).

III.III. The Models May Have Performed Individual Responses to the Pandemic

While epidemiological models have influenced individual behaviour through scientific advising and policy-making, they may have also been performative in a more direct way, by changing individual behaviour through information conveyed to the public. Individual behaviour during the pandemic was and is not simply determined by social-distancing and other public policy measures. Rather, how and whether social-distancing measures are taken up by individuals depends, to a significant extent, *both* on their willingness to comply with these measures *and* on how they interpret these measures.

Obviously, the effectiveness of public policy measures depends on the public's degree of compliance. For example, whether a policy restricting all social gatherings to the members of one household plus one additional person (a restriction in place in Germany through most of the pandemic) can contribute to a reduction in the reproduction rate R_0 of the virus depends on the extent to which people abide by this restriction. But the success of many public policy measures goes beyond the question of mere compliance. Many measures, such as many types of restrictions on social interactions, are compatible with a wide range of different social networks, including some quite extensive networks, and with a wide range of different types of spatial contacts among individuals. How such measures contribute to the slowing of the spread of the virus thus partially depends on how individuals interpret a given policy and how they implement it in adjusting their interactions. There are many different ways, for example, of following the 'one-household-plus-one-additional-person' rule: I might meet with a large number of different friends individually night after night, who all have their own largely non-overlapping group of friends, with whom they also meet; or I might radically cut down on my social contacts and form an isolated 'pod' with just one other person. Both behaviours comply with the rule, but the resulting social networks end up looking quite different. This is also the case in situations where explicit policy guidance is lacking—for example, the number of

infections that result from interactions in the workplace (which is one of the types of interaction explicitly modelled in the ICL model) will depend to some extent on how individuals adjust their behaviour in response to the model's predictions.

But to what extent individuals are willing to comply with policy measures, and how measures are interpreted and implemented in individual behaviour, depends on how serious individuals take the threat presented by the pandemic to be. This in turn depends, at least partly, on the information the public receives concerning the pandemic—information that often prominently includes model projections in infection, reproduction, and fatality rates (for example, Boseley 2020). For instance, when people are provided with this information through some kind of graphical representation, such as the graph in Figure 1, depicting how the demand for critical care beds could greatly exceed capacity, this might make the urgency of the situation particularly salient, and it might lead to different reactions than if this information had been conveyed through other means.

Thus, the predictions of epidemiological models and how these are understood by the public can play an important causal role in changing properties of the pandemic by influencing individual responses to the pandemic—a connection that is further underscored by the empirically well-supported assumption that people in many places reduced their contacts in advance of the implementation of social distancing policies (for example, Friedson et al. 2020; Sears et al. 2020). That is, our modelling efforts may interact with the very quantities and properties we are trying to model and can thereby change them: epidemiological models and the quantities modelled by them exhibit what Ian Hacking has called a "looping effect" (1999, 105).

For example, CovidSim contains a parameter representing the 'relative spatial contact rate given social distancing'—a parameter representing the effectiveness of social distancing measures, which, as Edeling et al. (2021) have shown, is particularly influential in determining the model's projections. To what extent the parameter value chosen provides an adequate representation of the effect of social distancing measures depends on how the measures are implemented in individuals' behaviour, which may be influenced by the model's projections.

Simpler models might not distinguish the institution of a policy from compliance with it, and simply model how universal compliance would affect infection rates. Yet in this case, too, the extent to which such a

model can provide adequate predictions concerning the effect of policy measures will depend on the degree to which the public abides with these measures. And the effectiveness of social distancing measures depends on how the model's results are communicated to and understood by the public.

IV. PERFORMATIVITY IN MODEL EVALUATION AND APPRAISAL

With these three potentially performative aspects of epidemiological models established, we are now in a position to draw out two ways in which recognizing performativity can assist us in the evaluation of these models. We consider, first, how performative aspects may compound the difficult task of successfully forecasting the course of an epidemic and, second, how they might be assessed in light of their ability to change the course of an epidemic.

Epidemiological models serve dual purposes: apart from their epistemic purpose of forecasting the course of an epidemic, they also serve the practical purpose of informing and guiding policy-making. These epistemic and practical purposes go hand in hand: on the basis of forecasts, policy-makers can choose policies that are likely to prevent unwanted outcomes. However, as we have seen in the previous section, the performativity of models complicates this task, because the implementation of these policies can steer the course of the epidemic away from the models' reference forecasts and can influence the adequacy of assumptions concerning the effectiveness of policy measures. If a modelling study does not include a range of scenarios reflecting different combinations of policies (some of which resemble those that will actually be implemented), some of the estimated parameters may turn out to be inaccurate due to the effects of the implemented policies (for example, levels of mobility), and the forecasts derived from the model may thus be inaccurate, too.

To be sure, many of the models that have been used to forecast the course of the Covid-19 pandemic, including the ones we have focused on here (Ferguson et al. 2020; Walker et al. 2020), do include a range of scenarios which project the impact of different (combinations of) policies. This can generate projections that better resemble the actual policy trajectories followed in a country, and models' predictive accuracy can accordingly be assessed by comparing these projections with observed outcomes (van Basshuysen and White 2021, forthcoming). However, gauging the accuracy of even these more pertinent scenario forecasts can face significant limitations, due to the fact that most models abstract from

endogenous behavioural responses, such as when individuals engage in social distancing even before government-issued directives take effect, adopt more rigorous hygiene and personal protection protocols, cancel travel plans, and so on. Such behavioural responses are poorly understood and difficult to anticipate (Friedman et al. 2021) and are thus difficult to incorporate into models (which is perhaps why extant models make no attempts to include them). *Report 9*, for instance, forecasts epidemic trajectories conditional on combinations of policies, but not on these types of spontaneous behavioural responses, potentially impairing the accuracy of important parameter values.[7] Furthermore, even sincere efforts at estimating these parameters may be limited, as behavioural responses can change over time (such as when individuals become tired of social distancing directives and depart from recommended behaviours, or when the public perception of the trustworthiness of expert and model projections changes). Even where modellers endeavour to conscientiously estimate crucial empirical parameters, and to re-parameterize models on an iterative basis (as is done for some of the models we discuss), it might not be possible to completely overcome these epistemic difficulties. The performativity of models thus presents a serious challenge to their predictive capacities.

However, while the performative aspects of epidemiological models can hamper their ability to successfully forecast the course of an epidemic, recognizing the performative dimension of these models can allow us to judge whether they have been successful in a second way, namely, by changing the course of an epidemic. That is, performativity may figure as an important criterion for appraisal beyond predictive accuracy.

When models are performative, their dual roles as epistemic instruments (that predict the evolution of a pandemic) and as practical instruments (that inform policy-making through reference and scenario forecasts) can make it difficult to evaluate them with respect to each purpose individually. As we have discussed, the performative effects of these

---

[7] It should be noted that not all distortions due to spontaneous behavioural changes are the effect of performativity being in play. A distinction might be drawn between individuals adapting their behaviour in response to model predictions, or in response to other kinds of information, such as current prevalence, which might be gathered and conveyed independently. It can be difficult, however, to determine which mechanisms are in play in a given case, as it might not be clear whether individuals' sources of information include model results. For our point that performativity may exacerbate the task of forecasting, it suffices that individual behaviour is *sometimes* a response to model results, which should be uncontroversial, given that these results and their graphical representations have figured prominently in public perception (as outlined above).

models on policy and individual behaviour can drive outcomes away from the models' predictions. When the predictions of models are undermined, moreover, this seems to undermine their epistemic credentials, making them unsuitable as a basis for public policy or individual behavioural response. When we focus only on the predictive capacities of such models, we seem to be in a bind here—the more they steer policy and behaviour, the less suitable they seem as a basis for steering policy and behaviour (unless they endogenize behavioural response—although, as we have just pointed out, this can be difficult).

One way out of this bind is to move model appraisal beyond focusing on models' predictive performance alone. We thus suggest that both predictive and performative capabilities should be considered side-by-side when appraising epidemiological models. This could proceed by evaluating some forms of performativity as a desirable attribute (and others as undesirable) in all-things-considered assessments of what models have done for us, epistemically and practically. Following Philip Kitcher (for example, 2011) in understanding and evaluating science through the roles it plays in, and partly *for*, society (where those roles often extend beyond purely epistemic ones), we suggest that a more comprehensive assessment of what epidemiological models have done for us (and whether they did a good job of it) should consider their performative aspects, too.

To use an analogy, consider a medical doctor telling her patient that he is likely to die of a heart attack within the next decade if he continues smoking, drinking, and doing little exercise. Such a prediction can be performative in helping the patient change his habits, allowing him to escape the unwanted prediction. Likewise, epidemiological models can have the performative ability to change important aspects of the phenomena they seek to capture. While this can adversely affect their predictive abilities, this does not necessarily imply that a model's suitability, adequacy, or usefulness is diminished, if we consider its performative impact to be a potential virtue. Just as we should not think that a doctor is unqualified because she cannot tell whether a patient would indeed have died of a heart attack counterfactually, or because she cannot accurately forecast how much longer, exactly, a patient would live under a changed exercise regime, we suggest that it is unhelpful to assess the utility of epidemiological models based on their predictive abilities alone. Epidemiological models have been a crucial resource for informing and justifying policy interventions, and may have contributed to shaping both the public's understanding of the pandemic and their behavioural response to it. The

ability to make such performative contributions can be understood as a desirable feature of these models, even if models' performativity can diminish their predictive prowess.

It is important to stress, however, that not all forms of model performativity are desirable. We might think, for instance, that some performative successes of epidemiological models can also have unwanted side-effects, such as when some members of the public begin to take the difference between observed outcomes (such as cumulative mortality) and widely publicized pessimistic reference scenarios as reason to call the severity or reality of the phenomena being modelled into question. One might also worry that by choosing particular outcome variables as key modelling targets, such as critical care demand or cumulative mortality, models have been performative in making these outcomes relatively more salient than various economic, societal, or psychological endpoints, and that models might thereby have contributed to steering policy and public response in ways that led to suboptimal societal outcomes. We do not take a stance on which performative aspects are desirable or undesirable here, but rather highlight that it is important to take both kinds of factors into account.

So how, exactly, are we to take models' performative aspects into account? Unfortunately, numerous complications arise when detailing the potential role of performativity in model appraisal. Should it play an evaluative role, providing a criterion by which the merits or deficiencies of models can be judged? Or should it be prescriptive, recommending the construction, selection, or use of models with certain performative capacities? The latter view seems highly problematic: we do not think that models can (or should) be made better by being made more performative, even if it were possible to agree that particular forms of performativity are socially desirable. Deceitfully meddling with a model in order to steer policy or behavioural response would raise significant concerns about illegitimate value-influences, and could severely undermine the epistemic credentials of models. Models are, and should remain, epistemic instruments in the first instance, and while some value-influences are rarely avoidable, model construction and use should proceed in epistemically responsible ways (see Elliott and McKaughan 2014; Parker and Winsberg 2018 for discussions of the roles of non-epistemic values in model construction and use).

To allay such concerns, we need principles that allow us to acknowledge models' performative aspects, good and bad, while

maintaining a commitment to the idea that models' epistemic functioning should remain untouched by considerations about their performativity. Unfortunately, formulating such principles, and providing a detailed account of how exactly we should weigh or integrate models' epistemic and performative aspects, involves various non-trivial questions that cannot be addressed here. These include: On which outcomes should we focus when gauging models' performative effects? How should we weigh considerations about models' epistemic and performative aspects? Should epistemic considerations enjoy lexical priority, or are there cases where performative virtues may compensate for epistemic shortcomings? Should all performative aspects be considered in model appraisal, or are there cases where we should distinguish between the model-as-such, the practice of model-use, and issues of science communication about model outputs (where some evaluations of performativity should perhaps pertain only to the latter two but not the former)? We suggest, for now, only the following provisional safeguard: performativity should only play an evaluative, but not a prescriptive, role in model appraisal. It may figure in judgments concerning whether the downstream performative aspects of models have been desirable or undesirable, but such judgments should not bear on decisions made at the stage of model construction, selection, or deployment—there should be no *wishful modelling*. Navigating the additional intricacies involved in formulating fine-grained principles is beyond the scope of this paper and must be addressed in future work.

Finally, we would like to highlight a rather different (and less problematic) route to taking models' performativity into account at the stage of model construction: instead of aiming to build models with specific performative goals in mind, we can aim to build models that endogenize their own performativity. Social scientists, as early as the 1950s, have made efforts to develop analytic machinery that can help anticipate the performative effects of public predictions (Grunberg and Modigliani 1954; Simon 1954). Despite significant difficulties, there has been progress with attempts to incorporate endogeneity into epidemiological models (see, for example, Avery et al 2020; Eksin, Paarporn, and Weitz 2019). Continuing these efforts could allow us not only to improve models' predictive capabilities, but also to answer criticisms that social scientists have been insufficiently involved in scientific advisory systems that inform policy. Social scientists could make an important contribution to the predictive endeavours undertaken by epidemiologists by helping

account for and anticipate the intricacies of the social and behavioural underpinnings that govern models' performativity.

V. CONCLUSIONS

We have argued that epidemiological models may have altered the course of the Covid-19 pandemic in three important respects: (i) through shaping the form of scientific policy advice, (ii) through shaping policy formation, and (iii) through influencing individual behavioural responses to the pandemic. Existing philosophical literature on epidemiological models and extant model evaluation studies have, however, focused primarily on predictive accuracy, and these performative aspects—and their implications—have been largely neglected. We have proposed that it may be important to take the performativity of models into consideration. This is, first, because it can exacerbate the difficulties involved in accurately predicting the course of a pandemic, and second, because certain types of performativity might be regarded as a desirable feature of epidemiological models. It thus might be fruitful to consider, in addition to predictive accuracy, the ways in which we can assess the success of models based partially on their performative aspects, and the role these might play in our fight against the pandemic.

REFERENCES

Avery, Christopher, William Bossert, Adam Clark, Glenn Ellison, and Sara Fischer Ellison. 2020. "An Economist's Guide to Epidemiology Models of Infectious Disease." *Journal of Economic Perspectives* 34 (4): 79–104.

van Basshuysen, Philippe, and Lucie White. 2021. "Bad Data and Flawed Models? Fact-Checking Winsberg et al.'s Case Against Lockdowns." *LSE Philosophy Blog*, January 26, 2021. https://www.lse.ac.uk/philosophy/blog/2021/01/26/bad-data-and-flawed-models/.

van Basshuysen, Philippe, and Lucie White. Forthcoming. "Were Lockdowns Justified? A Return to the Facts and Evidence." *Kennedy Institute of Ethics Journal.*

Birch, Jonathan. 2020. "Science and Policy in Extremis: The UK's Initial Response to COVID-19". Unpublished Manuscript. https://philpapers.org/archive/BIRSAP-4.pdf.

Boldyrev, Ivan, and Alexey Ushakov. 2016. "Adjusting the Model to Adjust the World: Constructive Mechanisms in Postwar General Equilibrium Theory." *Journal of Economic Methodology* 23 (1): 38–56.

Boseley, Sarah, 2020. "New Data, New Policy: Why UK's Coronavirus Strategy Changed." *The Guardian*, March 16, 2020. https://www.theguardian.com/world/2020/mar/16/new-data-new-policy-why-uks-coronavirus-strategy-has-changed.

Buck, Roger. 1963. "Reflexive Predictions." *Philosophy of Science* 30 (4): 359–369.

Callon, Michel. 1998. *The Laws of the Markets.* Malden, MA: Blackwell.

DHSC (Department of Health and Social Care). 2020. "Coronavirus Action Plan: A Guide to What You Can Expect Across the UK." Department of Health and Social Care, March 3, 2020. https://www.gov.uk/government/publications/coronavirus-action-plan/coronavirus-action-plan-a-guide-to-what-you-can-expect-across-the-uk.

Edeling, Wouter, Hamid Arabnejad, Robbie C. Sinclair, Diana Suleimenova, Krishnakumar Gopalakrishnan, Bartosz Bosak, Derek Groen, et al. 2021. "The Impact of Uncertainty on Predictions of the CovidSim Epidemiological Code." *Nature Computational Science* 1 (2): 128–135.

Eksin, Ceyhun, Keith Paarporn, and Joshua S. Weitz. 2019. "Systematic Biases in Disease Forecasting – The Role of Behavior Change." *Epidemics* 27: 96–105.

Elliott, K., and D. McKaughan. 2014. "Nonepistemic Values and the Multiple Goals of Science." *Philosophy of Science* 81 (1): 1–21.

Ferguson, Neil M., Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, et al. 2020. "Report 9: Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand." Imperial College COVID Response Team, March 16, 2020. https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/.

Friedman, Joseph, Patrick Liu, Christopher E. Troeger, Austin Carter, Robert C. Reiner JR, Ryan M. Barber, James Collins, et al. 2021. "Predictive Performance of International COVID-19 Mortality Forecasting Models." *Nature Communications* 12: 2609.

Friedson, Andrew I., Drew McNichols, Joseph J. Sabia, and Dhaval Dave. 2020. "Did California's Shelter-in-Place Order Work? Early Coronavirus-Related Public Health Effects." NBER Working Paper No. 26992. National Bureau of Economic Research, Cambridge, MA.

Grunberg, Emile, and Franco Modigliani. 1954. "The Predictability of Social Events." *Journal of Political Economy* 62 (6): 465–478.

Guala, Francesco. 2007. "How to Do Things with Experimental Economics." In *Do Economists Make Markets? On the Performativity of Economics*, edited by Donald MacKenzie, Fabian Muniesa, and Lucia Siu, 128–162. Princeton, NJ: Princeton University Press.

Hacking, Ian. 1999. *The Social Construction of What?* Cambridge, MA: Harvard University Press.

Henshel, Richard L. 1993. "Do Self-Fulfilling Prophecies Improve or Degrade Predictive Accuracy? How Sociology and Economics Can Both Be Right." *The Journal of Socio-Economics* 22 (2): 85–104.

Horner, Jack K., and John F. Symons. 2020. "Software Engineering Standards for Epidemiological Models." *History and Philosophy of the Life Sciences* 42 (4): 54.

Kitcher, Philip. 2011. *Science in a Democratic Society*. New York, NY: Prometheus.

MacKenzie, Donald. 2006. *An Engine, Not a Camera. How Financial Models Shape Markets*. Cambridge, MA: MIT Press.

MacKenzie, Donald, Fabian Muniesa, and Lucia Siu, eds. 2007. *Do Economists Make Markets? On the Performativity of Economics*. Princeton, NJ: Princeton University Press.

Mäki, Uskali. 2013. "Performativity: Saving Austin from Mackenzie." In *EPSA11 Perspectives and Foundational Problems in Philosophy of Science*, edited by Vassilios Karakostas and Dennis Dieks, 443–453. Dordrecht: Springer.

Nguyen, James, and Roman Frigg. 2021. "What are 'Scientific Models', and How Much Confidence Can We Place in Them?" *LSE Covid Blog*, February 10, 2021. https://blogs.lse.ac.uk/covid19/2021/02/10/what-are-scientific-models-and-how-much-confidence-can-we-place-in-them/.

Pahl, Hanno, and Jan Sparsam. 2016. "The IS–LMization of the General Theory and the Construction of Hydraulic Governability in Postwar Keynesian Macroeconomics." In *Enacting Dismal Science*, edited by Ivan Boldyrev and Ekaterina Svetlova, 151–181. New York, NY: Palgrave Macmillan.

Parker, Wendy S., and Eric Winsberg. 2018. "Values and Evidence: How Models Make a Difference." *European Journal for Philosophy of Science* 8 (1): 125–142.

Rhodes, Tim, Kari Lancaster, Shelley Lees, and Melissa Parker. 2020. "Modelling the Pandemic: Attuning Models to Their Contexts." *BMJ Global Health* 5 (6): e002914.

SAGE (Scientific Advisory Group for Emergencies). 2020a. "Potential Effect of Nonpharmaceutical Interventions (NPIs) on a COVID-19 Epidemic in the UK, 26 February 2020." Scientific Advisory Group for Emergencies, March 20, 2020. https://www.gov.uk/government/publications/potential-effect-of-non-pharmaceutical-interventions-npis-on-a-covid-19-epidemic-in-the-uk-26-february-2020.

SAGE (Scientific Advisory Group for Emergencies). 2020b. "Seventeenth SAGE Meeting on Covid-19, March 18." Scientific Advisory Group for Emergencies, May 29, 2020. https://www.gov.uk/government/publications/sage-minutes-coronavirus-covid-19-response-18-march-2020.

SAGE (Scientific Advisory Group for Emergencies). 2020c. "Summary of the Effectiveness and Harms of Different Pharmaceutical Interventions." Scientific Advisory Group for Emergencies, September 21, 2020. https://www.gov.uk/government/publications/summary-of-the-effectiveness-and-harms-of-different-non-pharmaceutical-interventions-16-september-2020.

Sayers, Freddie. 2020. "The Case for a Circuit-Breaker." *Unherd*, October 16, 2020. https://unherd.com/2020/10/the-case-for-a-circuit-breaker/.

Schroeder, S. Andrew. 2021. "How to Interpret Covid-19 Predictions: Reassessing the IHME's Model." *Philosophy of Medicine* 2 (1): 1–7.

Sears, James, Villas-Boas, J. Miguel, Villas-Boas, Vasco, and Villas-Boas, Sofia. 2020. "Are We #StayingHome to Flatten the Curve?" Department of Agricultural and Resource Economics, May 25, 2020; first version April 5, 2020; revised May, 2020. CUDARE Working Papers. Available at SSRN. https://ssrn.com/abstract=3569791.

Simon, Herbert A. 1954. "Bandwagon and Underdog Effects and the Possibility of Election Predictions." *The Public Opinion Quarterly* 18 (3): 245–253.

St. Onge, Peter, and Gaël Campan. 2020. "The Flawed COVID-19 Model That Locked Down Canada." Montreal Economic Institute, June 4, 2020. https://www.iedm.org/the-flawed-covid-19-model-that-locked-down-canada/.

Stewart, Heather, and Ian Sample. 2020. "How No 10's Relationship with its Scientists Broke Down." *The Guardian*, October 17, 2020. https://www.theguardian.com/world/2020/oct/16/scientists-call-to-arms-went-unheeded-how-the-relationship-with-no-10-broke-down.

Walker, Patrick, Charles Whittaker, Oliver Watson, Marc Baguelin, Kylie E C Ainslie, Sangheeta Bhatia, Samir Bhatt, et al. 2020. "Report 12: The Global Impact of COVID-19 and Strategies for Mitigation and Suppression." Imperial College COVID

Response Team, March 26, 2020. https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-12-global-impact-covid-19/.

Winsberg, Eric, Jason Brennan, and Chris W. Surprenant. 2020. "How Government Leaders Violated Their Epistemic Duties during the SARS-CoV-2 Crisis." *Kennedy Institute of Ethics Journal* 30 (3–4): 215–242.

Philippe van Basshuysen is a postdoctoral researcher at the Leibniz University Hannover and research associate at the Centre for the Philosophy of Natural and Social Science at the London School of Economics. His research is in the philosophy of economics and public policy, with a focus on public health.
Contact e-mail: <philippe.van.basshuysen@philos.uni-hannover.de>

Lucie White is a postdoctoral researcher at the Leibniz University Hannover, specializing in bioethics and the ethics of technology. She is currently working on a couple of projects which focus, in various ways, on policy responses to Covid-19.
Contact e-mail: <lucie.white@philos.uni-hannover.de>

Donal Khosrowi is a postdoctoral researcher at Leibniz University Hannover. His research interests and recent publications focus on causal inference in social science (especially in economics and evidence-based policy), scientific representation with models, and values in science.
Contact e-mail: <donal.khosrowi@philos.uni-hannover.de>

Mathias Frisch is a professor of philosophy at Leibniz University Hannover. His research interests lie in causation, climate modelling, and decision making under uncertainty.
Contact e-mail: <mathias.frisch@philos.uni-hannover.de>